Introduction
- New audio coding method based on 2D spectral recovery by convolutional neural network
  - Objective: better coding performance than current state-of-the-art codecs
- Novelty of 2D spectral recovery
  - Conventional method: spectral recovery on a block basis
  - High-band recovery based on low-band with optional high-band parameters
  - Proposed method: recovery of individual spectral coefficients based on neighboring information
    - Arrangement of transmitted data and recovered data in 2D check pattern

The proposed method can be integrated with USAC (state-of-the-art audio codec).

Performance
- Significantly better than USAC frequency-domain mode at 39.4 kbps

Proposed method: Structure of recovery
- Using previous frames to recover current frame
- Overall structure of 2D MDCT coefficients: $k_U \times 8$ MDCT magnitudes
  - Current frame and past 5 frames

Reconstruction of MDCT magnitudes using CNN
- $k_U$ is set to 608 (0~14.25 kHz at 48 kHz sampling rate).
- Determination of sign assignment of R
- Magnitude recovery of R
- Determination of sign transmission rule of R

Convolutional neural network (CNN)
- Activation function: Hidden layer: ReLU
- Output layer: tanh
- Optimizer: Adam
- Cost function: L1 cost function

Proposed method: Overall operation
- Quantization and coding of $Q_0$
  - Using quantizer and arithmetic coder of the USAC
  - Converting 2D MDCT coefficients into 1D data using two scanning patterns for entropy coding
- Proposed method integrated with USAC provides higher coding performance than USAC.